JOURNAL OF VIROLOGY, Aug. 1997, p. 5849–5860
0022-538X/97/$04.00+0

EXHIBIT B

Vol. 71, No. 8

Selection of a Nonconsensus Branch Point Is Influenced by an RNA Stem-Loop Structure and Is Important To Confer Stability to the Herpes Simplex Virus 2-Kilobase Latency-Associated Transcript

CLAUDE KRUMMENACHER,† JANICE M. ZABOLOTNY, AND NIGEL W. FRASER*

The Wistar Institute, Philadelphia, Pennsylvania 19104-4268

Received 27 February 1997/Accepted 25 April 1997

Herpes simplex virus type 1 latent infection in sensory neurons is characterized by the highly restricted transcription of viral genes. The latency-associated transcripts (LAT) family members are the only transcripts that can be identified in large amounts in latently infected cells. The most abundant LAT species is a 2-kb RNA that results from splicing of a rare primary transcript. Analysis of a LAT mutant virus (TB1) in cell culture revealed an aberrant splicing pattern and production of a stable small (0.95-kb) LAT intron. A panel of deletion constructs expressing truncated LAT in transiently transfected cells mapped the region influencing stability to the 3′ end of the LAT intron. This region encompasses the branch point and a putative stable stem-loop hairpin structure immediately upstream of the splice acceptor consensus polypyrimidine tract. Mutagenic analysis of the sequence in this region confirmed our hypothesis that the stem-loop structure is important for efficient splicing by influencing the selection of a nonconsensus branch point. Changes in this structure correlate with changes in branch point selection and production of an unstable 2-kb LAT.

Herpes simplex virus type 1 (HSV-1) is a human pathogen that causes lifelong infection punctuated with recurrent episodes of massive viral production (56). After an initial acute infection of the mucosa, HSV-1 infects nerve endings and travels to sensory ganglia, where it can establish a latent infection in neurons (38). Upon stress, the viral genome undergoes extensive transcription and replication, leading to the production of viral proteins and infectious particles (18, 51). In contrast to the initial acute infection or subsequent reactivation, the latent stage is characterized by a tight transcriptional repression of all classes of viral genes, with the exception of a diploid locus in the long repeat elements known as the latency-associated transcript (LAT) gene (9, 44, 52). The most abundant LAT species is a 2-kb nonpolyadenylated transcript referred to as the 2-kb LAT (44, 50, 52). Removal of a short intron in the 2-kb LAT leads to the production of a small variant of 1.5 kb (47). This smaller LAT is observed only during latency in neurons, whereas the 2-kb LAT is also expressed during productive infection with late gene kinetics (46, 55). The primary LAT transcript, known as the minor LAT species (mLAT), extends for 8.3 kb from the main LAT promoter to the first downstream polyadenylation signal (29; reviewed in reference 17).

The role of the LAT gene products during the life cycle of HSV is not clearly understood. Since they are expressed during the latent phase of infection, their role in the establishment and maintenance of and reactivation from latency has been examined. Some HSV-1 LAT-negative viruses display limited efficiency in establishing latency in neurons (41) and/or reduced reactivation kinetics in animal models (3, 5, 22, 26, 33, 53). The large number of overlapping genes (ICP0, ICP4, and γ-34.5) and other newly described transcripts (ORF P, ORF O, and LS/Ts) in this region often prevents a direct assignment of the observed phenotype to LAT; indeed, the functional LAT gene products are not precisely defined. Despite much effort (13), and in contrast to findings with bovine herpesvirus (42), no functional LAT-derived protein has been clearly identified. Since the 2-kb LAT accumulates to high levels in infected neurons, it is thought by many to be an active RNA. However, the mechanism of synthesis for this LAT transcript and its function remain elusive. Since it is partially complementary to the 3′ end of the ICP0 mRNA, the 2-kb LAT has been proposed to be involved in an antisense suppression mechanism (16, 52).

The 2-kb LAT, beginning some 600 bp downstream of the main LAT promoter (12, 59) and lacking most of the structural characteristics of an mRNA such as a cap or poly(A) tail (11, 54), has been postulated to be a stable intron accumulating after splicing of the minor LAT (mLAT) (12). The ability of the 2-kb LAT sequence to be excised from a β-galactosidase mRNA was first demonstrated experimentally by Farrell et al. (16). Recent structural data reporting the nonlinear nature of both the 1.5- and 2-kb LATs (37, 57) and mapping of the branch point of this RNA (58) confirmed the intronic origin of the 2-kb LAT.

In the course of natural RNA processing, introns are removed from mRNA by a succession of molecular events involving specialized factors such as small nuclear ribonucleoproteins (30). Proper recognition of splicing signals, such as splice donor and acceptor sites, and branch point sequences by small nuclear RNPs are essential steps in this process (19, 35). Consensus branch sites are efficiently selected, whereas nonconserved sites are poorly recognized and often lead to suboptimal alternative splicing (48) or exon skipping (21, 25, 36). In addition to the branch point itself, surrounding regions (7, 14), RNA secondary structures (15, 20, 43), and polypyrimidine tract sequences (32, 39) are known to influence branch point selection. Introns are usually rapidly degraded by a mul-

* Corresponding author. Mailing address: The Wistar Institute, 3601 Spruce St., Philadelphia, PA 19104-4268. Phone: (215) 898-3847. Fax: (215) 898-3849. E-mail: fraser@wista.wistar.upenn.edu.

† Present address: Department of Microbiology, School of Dental Medicine, University of Pennsylvania, Philadelphia, PA 19104-6002.

tistep process involving nucleases and specialized debranching enzymes (1, 30, 31). The 2-kb LAT intron has been reported to be more resistant to the action of debranching enzymes (37). Other stable introns of cellular (34) and viral (6) origin have been described. A general mechanism underlying intron resistance to degradation, however, has not been elucidated. Different strategies have evolved to confer stability to RNA, particularly to mRNA. These include (i) secondary structures serving as protein-docking sites, (ii) the presence of a long polyadenylated tail, and (iii) possible sequestration in a nuclease-free environment (40).

In this study we investigated the sequence requirements for the production of a stable 2-kb LAT in tissue culture. On the basis of observations from an HSV-1 LAT mutant producing a truncated LAT intron (4) and a panel of mutated LAT expression plasmids, we mapped a region at the 3′ end of the 2-kb LAT which is important for stability. The nucleotide sequence involved in the formation of a putative secondary structure in the tail of the 2-kb LAT lariat was found to be responsible for nonconventional branch point selection, influencing both splicing efficiency and stability of the intron.

MATERIALS AND METHODS

Cells and virus. Cos-1 cells were grown in Iscove's medium supplemented with 5% calf serum. CV-1 cells were maintained in Eagle's minimal essential medium plus 5% calf serum. HSV-1 mutant TB1 and HSV-1 F were prepared as previously described (4, 9, 10).

Plasmid construction. The 2.8-kb *Pst*I-*Mlu*I restriction fragment encompassing the LAT gene of HSV-1 F was subcloned into the *Eco*RI and *Hin*dIII sites of pcDNA3 (Invitrogen) as described previously (58). The resulting plasmid, pcDNA3Pst-Mlu, served as the backbone for subsequent mutations and will be referred as the wild-type (wt) plasmid. Mutants pΔHpa and pΔXcm were generated by removal of the corresponding restriction fragment and self-ligation (prior Klenow filling in was necessary in the case of pΔXcm). To construct pΔBfa, the 2.8-kb *Eco*RI-*Hin*dIII fragment from pcDNA3Pst-Mlu was isolated and digested with *Bfa*I. The resulting 2,110-bp *Hin*dIII-*Bfa*I and 400-bp *Bfa*I-*Eco*RI fragments were purified and ligated directly in the *Eco*RI-*Hin*dIII-digested wt vector. Plasmids pHλ+ and pHλ− were generated by the insertion in pΔHpa of the 440-bp *Hpa*I fragment consisting of lambda phage DNA from plasmid pNF1. In pHλ−, the phage DNA insert is in the same orientation as in the TB1 virus; it is in the opposite orientation in pHλ+.

The pBam clone was generated by site-directed mutagenesis (23) to create a unique *Bam*HI site toward the 3′ end of the 2-kb LAT. Fragments from pcDNA3Pst-Mlu were amplified by PCR with primers PFPML (GGGGCATC ACGTGGTTACCC) and PRBAM (GAGACAAGAGGAAGGATCCCTCG GC) or with PFBH1 (AGGGATCCTTCCTCTTGTCTCCCTCCCAGG) and PRP31 (GCCAGTGTGATGGATATCTGC). Both amplified products were purified from agarose gels and mixed equally in a second PCR with PFPML and PRP31 as amplification primers. The 700-bp final product was digested with *Eco*RI and *Bbr*PI (*Pml*I) and cloned into the corresponding sites of vector pcDNA3Pst-Mlu. Following the same procedure, pCons was generated with pCR primers PRPB10 (GGAAAGGTCAGTGGGCCCCCGCGCTGCTTCTGG) plus PFPML and PFPB9 (CACTGACCTTTCCTCTTGTCTCCCTCCCAGG) plus PRP31 in the first-step PCR amplifications.

Mutant pY+ was generated by the insertion of a linker (GATCGTACTAAC) containing the yeast consensus branch point sequence into the *Bam*HI site of pBam. In pY−, the same linker was inserted in the opposite orientation (GAT CGTTAGTAC). Plasmid pΔPB was obtained after removal of the *Pml*I-*Bam*HI fragment from pBam, Klenow filling in, and self-ligation of the plasmid. To generate plasmids pΔ1, pΔ2, and pΔ3, DNA from wt vector was amplified by PCR with, respectively, PFPB1 (CCACACGTGAGACCCCCGAGATGGGC AGG), PFPB2 (CGCCACGTGGGACGGCCCCGGAAGTCTCC), PFPB3 (TT CCACGTGCCCCCGCGGCCCAGAAGCAGC), and PRP31; the fragments were gel purified, digested with *Eco*RI and *Pml*I, and cloned into the corresponding sites of pcDNA3Pst-Mlu. To obtain pΔ4, DNA from pcDNA3Pst-Mlu was amplified with PRPB6 (GGAAACCTCCCTCCCGAGGAAGTGTGCCCGGAAG ACG) and PFPML or with PFPB7 (GAGGGAGGTTTCCTCTTGTCTCCCTC CCAGG) and PRP31; amplified fragments were gel purified and equally mixed in a PCR with PFPML and PRP31 as primers. The final product was digested with *Pml*I and *Eco*RI and cloned into the corresponding sites of the wt vector. Mutant pΔ5 was generated by the same method but with PRPB6 replaced by PRPB8 (GGAAACCTCCCTCCCGCGGGCGGCCTCACGCGCTACC). Mutant pΔG was engineered in the same way with primers PFDBG (CGTGGCG CGTCTTACACTTCCTCGGCCC) and PRP31 or primers PFPML and PRDBC (AAGACGCGCCACGCGGAGGC). The resulting products were gel purified and mixed in a second amplification reaction with primers PFPML and PRP31. The final 700-bp product was purified, digested with *Pml*I and *Eco*RI, and cloned in the corresponding sites of pΔ3. A similar procedure was used for pΔA with primers pFDELA (CAGCGCCGGGGGCCTTTCCTCTTGTCTCCCTCCCAGG) and PRDELT (AAAGGCCCCGGCGCTGCTTCTGG) in the first-step PCRs.

All amplification reactions were performed under the same conditions: 35 cycles of 1 min at 94°C, 1 min at 60°C (58°C), and 1 min at 72°C in 1× PCR buffer [60 mM Tris-HCl (pH 9.5), 15 mM $(NH_4)_2SO_4$, 1.5 to 2.0 mM $MgCl_2$]–0.2 mM deoxynucleoside triphosphates–1 μM each oligonucleotide–0.5 U of *Taq* polymerase (Fisher Scientific). Five cycles with an annealing temperature of 58°C were performed in the absence of oligonucleostides prior to standard amplification in the second-step PCR when two PCR fragments had to be linked as the template.

Transfection and RNA extraction. A 20-μg portion of plasmid per 10-cm dish was transfected into subconfluent Cos-1 cell monolayers by calcium phosphate precipitation (2). The cells were left with the precipitate for 16 h, shocked in phosphate-buffered saline–15% glycerol for 2 min, washed with phosphate-buffered saline, and incubated for 22 to 24 h at 37°C. Alternatively, subconfluent CV-1 cells were infected with 1 PFU of TB1 or HSV-1 F virus per cell. RNA was isolated 16 h postinfection. Transfected or infected cells were lysed in 3.5 ml of lysis buffer (4 M guanidine isothiocyanate, 0.5% sodium *N*-lauroylsarcosine, 100 mM β-mercaptoethanol, 25 mM sodium citrate [pH 7.0], 0.1% antifoam A). DNA was sheared for 10 s by a mechanical disrupter (Brinkmann Instrument Inc.). Total RNA was pelleted through a cushion of 5.7 M CsCl–0.1 M EDTA by centrifugation at 150,000 × *g* for 20 h at 18°C. The RNA pellet was resuspended in H_2O and precipitated with ethanol. RNA was stored in ethanol at −70°C.

Northern blot analysis. RNA was resuspended in H_2O, and the absorbance at 260 nm was measured spectrophotometrically. A 5-μg portion of total RNA was treated with glyoxyal, separated on a 1.2% agarose gel, and transferred on GeneScreen Plus (NEN) membrane as previously described (45). The filters were prehybridized for 2 h at 50°C in 50% formamide–10% dextran sulfate–1× Denhardt's solution–1% sodium dodecyl sulfate–5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–1 mM EDTA–0.1% denatured salmon sperm DNA. The ^{32}P-labeled nick-translated probe was heat denatured, added to the prehybridization mix, and incubated overnight at 50°C. The blots were washed in decreasing concentrations of SSPE (1×, 0.5×, and 0.1×) (1× SSPE is 0.18 M NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA [pH 7.7])–1% SDS at 65°C, twice for 20 min per wash, and exposed on a PhosphorImager screen (Molecular Dynamics). The bands were quantified with ImageQuant version 1.1 software.

RT-PCR. To characterize the splice junctions of LAT, cDNA was synthesized from 1 μg of DNase-treated total RNA with the Superscript II preamplification kit (Gibco BRL). The manufacturer's protocol for high GC content RNA was followed with the use of both poly(dT) and random hexanucleotides as primers. PCR was performed on 20 ng of cDNA with 1 μM primers exon1 (GCTCCAT CGCCTTTCCTGTT) and exon2N (TCCTCTGCCTCTTCCTCCTCG) in 1× PCR buffer–0.2 mM deoxynucleoside triphosphates–0.5 U of *Taq* polymerase (Fisher Scientific) in a 50-μl reaction volume. The conditions of the 35 amplification cycles were 1 min at 94°C, 1 min at 60°C, and 1 min at 72°C. A 5-μl sample of the reaction product was electrophoresed on a 2% agarose gel and visualized by ethidium bromide staining. Amplified bands were extracted from the gel with a Geneclean kit (Bio 101, Inc.), cloned into a TA vector (Invitrogen), and sequenced with an ABI Prism cycle sequencer.

To map the branch point of the different introns, reverse transcription-PCR was performed as described previously (58). Briefly, LAT-specific cDNA was synthesized from 10 μg of DNase-treated total RNA with the "branch" oligonucleotide (AGAAGCAGGTGTCTAACCTACNNN) as primer. This cDNA was subjected to a PCR amplification performed with primers PFPML (GGGG CATCACGTGGTTACCC) and PREND (AGAAGCAGGTGTCTAACCTA C). The PCR products were isolated from 2% agarose gels, cloned into a TA vector (Invitrogen), and sequenced.

RESULTS

Production of altered LAT after insertion of a lambda sequence in the 2-kb LAT. The HSV-1 insertion mutant TB1, which has a 168-nucleotide deletion in the middle of the 2-kb LAT region replaced by 440 nucleotides of lambda phage DNA, has been shown to be unable to express a stable 2-kb LAT in lytically infected cells or latently infected animals (4). To determine the mechanism underlying this particular phenotype, a LAT-expressing plasmid vector was constructed in which the TB1 mutation was re-created. In the parental vector, a 2.8-kb *Pst*I-*Mlu*I fragment encompassing the 2-kb LAT of HSV-1 F is expressed under the control of a cytomegalovirus immediate-early promoter. The bovine growth hormone polyadenylation signal is present downstream of the HSV-1 insert, ensuring proper termination of transcription. We have previously identified the three transcripts expressed by this plasmid as a truncated mLAT primary transcript (3.4 kb), the 2-kb LAT